Security Framework for Quantum Distance-Bounding

Kevin Bogner[1], Aysajan Abidin[1], Dave Singelee[2], Bart Preneel[1]

[1]COSIC, KU Leuven, Leuven, Belgium
[2]DistriNet, KU Leuven, Leuven, Belgium
kevin.bogner@kuleuven.be, aysajan.abidin@kuleuven.be,
dave.singelee@kuleuven.be, bart.preneel@kuleuven.be

Abstract

Distance-bounding (DB) protocols let a verifier upper-bound a prover's physical distance by timing rapid challenge-response exchanges. Quantum communication promises simpler DB protocols with stronger security guarantees, yet existing quantum distance-bounding (QDB) proposals are analysed in ad-hoc models and, to the best of our knowledge, lack a common game-based treatment of standard fraud attacks. We contribute (i) a reusable security framework for QDB that fixes system and timing assumptions, specifies a quantum-capable adversary model, formalises distance-, mafia-, and terrorist-fraud experiments, and includes a simple i.i.d. depolarizing noise model; and (ii) an application of this framework to a published QDB protocol. For this protocol we characterise the honest per-round acceptance probability under noise and lift it to the multi-round setting, yielding explicit completeness guarantees as a function of the number of fast rounds, the acceptance threshold, and the noise parameter. For active adversaries we bound the per-round success probability of distance-fraud attacks and analyse the best known mafia-fraud strategy, deriving corresponding multi-round soundness bounds. We also show that the protocol is inherently insecure against terrorist-fraud in our model. The framework cleanly separates protocol-independent definitions from protocol-specific analysis and can be used to evaluate existing and future QDB protocols on a common basis.

Keywords: Quantum distance-bounding, Distance-bounding, Provable security, Quantum cryptography, Quantum communication

1 Introduction

Verifying that a remote device is within a claimed physical distance is a fundamental primitive for access control, secure ranging, and secure localisation. *Distance-bounding* (DB) protocols achieve this by measuring the round-trip time of challenge-response exchanges, allowing the verifier to upper-bound its physical distance to the prover [1]. Typical applications include contactless authentication (e.g. RFID and NFC tokens), keyless entry systems, and secure ranging and localisation in wireless networks [2]. Classical DB designs resist both *distance-fraud* (DF), where a far-away dishonest prover cheats on its location, and *mafia-fraud* (MF), a relay attack by two external adversaries. The strongest threat is *terrorist-fraud* (TF), where a dishonest prover actively collaborates with a nearby helper. Classical countermeasures against TF remain expensive in terms of additional hardware or bandwidth, especially for low-cost devices [3].

At a high level, a DB protocol proceeds in three phases. In a *slow* (initialisation) phase, which is not time-critical, the verifier and prover exchange nonces, derive session-specific secrets from long-term keys, and set up all parameters needed for the distance test. This is followed by a *fast* phase consisting of n time-critical rounds: in each round the verifier sends a fresh challenge, the prover must respond immediately, and the verifier measures the round-trip time to infer an upper bound on the distance. Finally, in the *decision* phase, the verifier checks

that each accepted round is both timely and logically consistent with the protocol rules, and accepts the session if the number of accepted rounds is at least a threshold τ. We will use this terminology (slow / fast / decision phase) for both classical and quantum distance-bounding protocols throughout the paper.

Quantum communication promises simpler and stronger DB because non-orthogonal quantum states cannot be perfectly copied [4], preventing relay adversaries from pre-computing answers. Several *quantum distance-bounding* (QDB) protocols have been proposed in recent years, starting with the protocol of Abidin *et al.* [5] and its follow-ups [6, 7, 8, 9]. On the classical side, DB protocols have been studied in a unified, game-based framework that fixes system assumptions, threat experiments, and soundness notions [10]. In contrast, existing QDB proposals are analysed in protocol-specific models and, to the best of our knowledge, none comes with a formal security proof against standard distance-, mafia-, and terrorist-fraud attacks. We aim to close this gap for QDB in this paper.

In this work we take a step towards such a foundation. We isolate a minimal set of system and timing assumptions for QDB, specify a quantum-capable adversary model, formalise DF/MF/TF experiments in this setting, and connect per-round cheating probabilities to multi-round security. We then apply this framework to a concrete QDB protocol [6] and obtain explicit completeness and soundness guarantees as functions of the number of fast rounds n, the threshold τ, and a simple noise parameter.

For context, we briefly compare the analysed protocol with the classical Hancke-Kuhn style DB protocol as a baseline [2].

The protocol we analyse [6] follows the same high-level structure as the classical Hancke-Kuhn DB protocol [2], consisting of a slow setup phase, n time-critical fast rounds, and a final decision that accepts iff at least τ rounds satisfy the time-of-flight bound and the protocol's value check. The key difference lies in the fast-phase workload and the resulting system assumptions. Classical Hancke-Kuhn-style DB exchanges classical challenge and response bits and therefore requires primarily tight latency engineering on a classical channel. In contrast, QDB replaces the fast-phase bit exchange by single-qubit transmissions, which introduces the cost of quantum state preparation and measurement (and a quantum channel), while leaving the slow phase and its symmetric-key structure essentially unchanged. Moreover, completeness and parameter sizing must account for quantum noise on the fast-phase transmissions (captured by our noise model in Section 4.4). From a security perspective, both settings rely on time-of-flight, fresh challenges, and symmetric-key assumptions in the slow phase, but QDB additionally exploits quantum-mechanical constraints such as no-cloning and measurement disturbance [4]. In our setting, the per-round secrets are derived via a *quantum-secure* PRF and we model a quantum-capable (QPT) adversary controlling both the classical and quantum channels (see Sections 4 and 5).

For a simple quantitative comparison, we use Hancke-Kuhn [2] as the closest classical baseline, since both protocols have the same slow/fast/decision structure and a lightweight symmetric-key slow phase. To keep this discussion simple at this stage, Table 1 fixes an idealised noiseless setting with strict threshold $\tau = n$. The table reports the smallest number of rounds needed to reach a session false-accept probability at most 2^{-80}. The full threshold/noise trade-off is analysed later in Section 6.

The comparison shows that the present QDB protocol [6] should not be viewed as a drop-in replacement for classical DB. While QDB improves the per-round distance-fraud bound, its best known mafia-fraud success probability is less favourable and the fast phase requires quantum hardware. Still, QDB remains worth studying. Quantum communication changes the fast-phase attack surface, and earlier QDB work already argued that early-detect/late-commit strategies do not transfer directly, since an adversary that does not know the correct basis cannot measure the challenge without disturbing the state [5]. At the same time, QDB introduces its own implementation assumptions and attack surface, as illustrated by photon-number-splitting attacks; indeed, the protocol of Abidin [6] was proposed as an improvement over the 2016 variant

Table 1: Comparison with the classical Hancke-Kuhn [2] baseline in an idealised noiseless setting with strict threshold $\tau = n$. For Hancke-Kuhn, the standard 3/4 DF/MF per-round baseline is summarised in [3]. P_{FA} denotes the false-acceptance probability. The QDB mafia-fraud row uses the best known mafia-fraud attack [11].

Metric	Hancke-Kuhn DB [2]	Abidin 2019 QDB [6]
Fast-phase transmission	1 classical bit each way	1 qubit each way
Per-round DF	3/4	1/2
Per-round MF	3/4	7/8
Rounds for DF, $P_{\mathrm{FA}} \leq 2^{-80}$	193	80
Rounds for MF, $P_{\mathrm{FA}} \leq 2^{-80}$	193	416

in this respect, and later entanglement-based proposals continue to explore this design space [7]. We therefore do not claim that the present QDB protocol already dominates classical DB on all metrics; rather, QDB offers additional physical constraints and a different design space that may enable stronger future protocols, which motivates a common evaluation framework.

To make the guarantees of our framework interpretable and comparable across protocols, we state upfront the system assumptions and modelling choices on which the main results rely. These assumptions are formalised in Sections 4 and 5 and are used consistently throughout the protocol analysis in Section 6.

Timing model. As in classical DB, our security guarantees rely on an explicit time-of-flight constraint. In each fast round, the verifier timestamps the moment a challenge state is emitted and accepts the round only if the corresponding response is received within a strict deadline derived from the distance bound (see Section 4.3). Any fixed device latencies, such as state preparation and measurement delays, are assumed to be either negligible or accounted for by calibration and absorbed into the effective distance bound. This enforcement is necessary as it implies that any response generated entirely outside the verifier's proximity and received by the deadline cannot causally depend on the fresh challenge of that round. This property is a direct consequence of no-superluminal signalling (formalised as Lemma 4).

Adversary model and channel control. We consider a quantum-capable adversary that controls both the classical and quantum channels, subject only to the constraints of quantum mechanics and relativity. The precise capabilities used in the DF/MF/TF experiments are formalised later in Section 5.

Noise assumptions. To capture imperfect quantum communication in *honest* executions, we adopt a simple, protocol-agnostic noise model for the quantum channel (see Section 4.4). This model is used only for analysing honest executions and parameter sizing. In contrast, our security bounds are derived against an *idealised noiseless adversary* with perfect channels and measurements. This is a conservative choice, as any physical imperfections affecting the attacker can only reduce its ability to return correct, timely responses under the same decision rule.

Threat scope. We focus on the standard fraud threats from classical DB, namely distance fraud, mafia fraud, and terrorist fraud, and formalise them via game-based experiments (Section 5). Our analysis is protocol-level and does not aim to model implementation-specific side channels or physical-layer attacks such as multi-photon effects, detector blinding, device fault attacks, or adversarially induced losses, nor do we treat alternative threat variants such as distance hijacking. These aspects are orthogonal to the core framework and are best addressed by extending the system model.

Our contributions are as follows.

- We provide a reusable, game-based security framework for proving the security of QDB protocols, covering system and timing assumptions, quantum adversary models, soundness definitions, and a simple noise model.

- We apply this framework to an already proposed QDB protocol [6] and derive explicit bounds on completeness and on DF/MF soundness in the multi-round setting.

The remainder of the paper is organised as follows. In Section 2 we position our contribution within the existing literature on classical DB, QDB, and related quantum position verification (QPV). In Section 3 we introduce the notation for quantum states and the probabilistic tools used later. Section 4 formalises QDB protocols, and Section 5 introduces the threat model and security experiments. Section 6 revisits the QDB protocol of Abidin [6] and applies the framework to it. Section 7 concludes the paper.

2 Related work

DB was introduced by Brands and Chaum as a cryptographic primitive for upper-bounding the physical distance to a prover through time-of-flight measurements in rapid challenge-response rounds [1]. Since then, classical DB has produced lightweight designs for constrained devices, most notably Hancke-Kuhn-style protocols, and a large body of work on standard fraud threats, namely DF, MF, and TF, under tight latency and noisy-channel constraints [2, 3]. On the formal side, classical DB has also benefited from reusable game-based security frameworks that make the system assumptions explicit and define DF/MF/TF experiments in a uniform way [10]. Our QDB framework is derived from this line of work and adapts it to the quantum setting.

Within this landscape, QDB replaces the classical fast-phase bit exchange by quantum communication, typically to exploit no-cloning and measurement disturbance as additional constraints on adversaries [4]. Early work by Abidin *et al.* studied the feasibility of this approach in Hancke-Kuhn-style protocols using BB84-type states and discussed timing, hardware-delay considerations, and informal DF/MF/TF success estimates [5]. Abidin later refined this line with an updated QDB protocol and a discussion of implementation-driven attacks such as photon-number splitting [6]. A related hybrid design by Abidin revisited an earlier qubit-based relay-attack-detection scheme and proposed an improved timed protocol with classical-bit challenges and qubit responses [12]. More recent proposals extend the design space to entanglement-based, mutual, and continuous-variable variants [7, 8, 9].

Despite this progress on protocol design, QDB security analyses remain largely protocol-specific. Existing works typically adopt bespoke system models and analyse selected attack strategies, rather than a common game-based treatment of DF/MF/TF with a standard lifting from per-round to multi-round security. Verschoor's re-analysis illustrates this gap by identifying stronger mafia-fraud strategies for published QDB protocols and by showing that TF resistance can depend subtly on the exact experiment definition [11]. To the best of our knowledge, QDB still lacks a unified treatment analogous to the classical DB frameworks [10].

A related, but distinct, line of work is *quantum position verification* (QPV). QPV uses a party's geographical position as the credential and typically involves multiple verifiers under relativistic constraints, whereas QDB studies shared-key proximity authentication between a verifier and a prover. We therefore do not adopt QPV security notions directly; we only note that, as in QPV, security depends critically on making the timing assumptions and adversary resources explicit [13].

3 Preliminaries

In this section we collect the necessary preliminaries for the security analysis. We start by introducing the notation for quantum states and measurements, which are inspired by the BB84 QKD protocol [14], and conclude with a concentration bound that we repeatedly use in later sections. The QDB protocol [6] that we will apply the framework to is recalled in Section 6.1.

3.1 Notation for quantum states and measurements

Let $a \in \{0,1\}$ index the basis ($a = 0$ is the computational Z basis, $a = 1$ is the diagonal X basis). Define the four BB84 states

$$|0\rangle_0 = |0\rangle\,, \quad |1\rangle_0 = |1\rangle\,, \qquad |0\rangle_1 = |+\rangle\,, \quad |1\rangle_1 = |-\rangle\,,$$

where

$$|\pm\rangle = \tfrac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle).$$

For $r \in \{0,1\}$ we write the rank-one projectors as

$$\Pi_{a,r} := |r\rangle_a \langle r|_a\,.$$

We write $\langle r|_a := (|r\rangle_a)^\dagger$ for the associated bra. We also write $\Pi_\psi := |\psi\rangle\langle\psi|$ for the projector onto a pure state $|\psi\rangle$. Measuring $|\psi\rangle$ in basis a yields outcome r with probability

$$\Pr[r] = \left\|\Pi_{a,r}\,|\psi\rangle\right\|^2 = \langle\psi|\,\Pi_{a,r}\,|\psi\rangle\,.$$

3.2 Concentration bound

We use a single tail bound for our QDB protocol, to bound the success probability of the protocol in an honest or malicious setting. The key requirement is *adaptivity-robustness*: even if the adversary chooses its action in round i *after seeing everything that happened so far*, the bound still applies as long as that round's conditional success probability is capped. For $u, v \in (0,1)$, let

$$D(u\|v) := u\ln\frac{u}{v} + (1-u)\ln\frac{1-u}{1-v}$$

denote the binary relative entropy between Bernoulli(u) and Bernoulli(v). In our applications, u will be the threshold ratio τ/n (number of required accepted rounds τ divided by the total number of rounds n), and v will be a bound p on the per-round success probability.

This is a Chernoff-type bound for adaptively generated Bernoulli variables: the indicators I_i may be chosen based on the full history $\mathcal{F}_{i-1}$, and we only require the conditional success probability $\mathbb{E}[I_i \mid \mathcal{F}_{i-1}]$ to be bounded by p. In the i.i.d. case it reduces to the standard Chernoff bound.

Lemma 1 (Chernoff bound under adaptivity [15])**.** *Fix any process of fast rounds and let $(\mathcal{F}_i)_{i=0}^n$ be the filtration where $\mathcal{F}_i$ denotes the information available* after round i has completed *(equivalently: just before round $i+1$ begins), with $\mathcal{F}_0$ the information at the start of the fast phase. Concretely, $\mathcal{F}_i$ contains*

- *the* public transcript *up to round i (all classical messages, timing/acceptance flags of rounds $1,\dots,i$, and any randomness already revealed); and*
- *the adversary's* internal state *(its private classical registers and any quantum registers it keeps between rounds, including entanglement with systems not yet measured).*

Let $I_i \in \{0,1\}$ be the indicator that round i is accepted, and let $S = \sum_{i=1}^n I_i$. If the single-round conditional success is capped as

$$\mathbb{E}[\,I_i \mid \mathcal{F}_{i-1}\,] \le p \qquad \text{for every } i \in [n],$$

then for any $\tau \in (np, n]$,

$$\Pr[S \ge \tau] \;\le\; \exp\Big(-n\,D\big(\tfrac{\tau}{n}\,\|\,p\big)\Big) \quad \textit{(upper tail).}$$

Similarly, if $\mathbb{E}[\,I_i \mid \mathcal{F}_{i-1}\,] \ge p$ for all i, then for any $\tau \in [0, np)$,

$$\Pr[S \le \tau] \;\le\; \exp\Big(-n\,D\big(\tfrac{\tau}{n}\,\|\,p\big)\Big) \quad \textit{(lower tail).}$$

4 Model for Quantum Distance-Bounding

Our model adapts the classical DB framework of Boureanu et al. [10] to the quantum setting; the definitions and security experiments introduced here and in Section 5 follow their structure closely.

4.1 Complexity

Let $\lambda \in \mathbb{N}$ denote the security parameter, which controls the security-efficiency tradeoff: as λ grows, costs may increase while any adversary's success probability drops to a negligible function in λ. Throughout this paper,

- a *probabilistic polynomial-time* (PPT) algorithm is a classical probabilistic Turing machine whose running time is bounded by a polynomial in the security parameter λ (and the length of its explicit inputs);
- a *quantum polynomial-time* (QPT) algorithm is a quantum algorithm whose running time is bounded by a polynomial in λ (and the input length);
- $\mathrm{poly}(\cdot)$ denotes an unspecified polynomial; and
- a function $\mathrm{negl}(\cdot)$ is *negligible* if for every polynomial $p(\cdot)$, $\mathrm{negl}(\lambda) \leq 1/p(\lambda)$ for all sufficiently large λ.

All honest-party algorithms are efficient; classical algorithms are PPT and quantum algorithms are QPT.

Let $n \in \mathbb{N}$ be the number of time-critical challenge-response rounds in the *fast phase* of one protocol execution; we treat n as a tunable parameter.

4.2 Participants

A verifier $\mathcal{V}$ and a prover $\mathcal{P}$ interact over a channel fully controlled by a quantum-capable adversary $\mathcal{A}$.

Definition 1 (Quantum Distance-Bounding protocol [10])**.** A QDB protocol is a tuple

$$\mathrm{QDB} \;=\; (\mathsf{KeyGen}, \mathcal{P}, \mathcal{V}, B),$$

where

1. $\mathsf{KeyGen}(1^\lambda) \to x$ is an efficient (classical or quantum) algorithm that outputs a classical key $x \in \{0,1\}^\lambda$;
2. $\mathcal{P}(x)$ is an interactive QPT algorithm;
3. $\mathcal{V}$ is an interactive QPT algorithm. We write $\mathcal{V}(x, r) \to \mathrm{Out}_\mathcal{V} \in \{0,1\}$ for its execution on key x with internal randomness r, where $\mathrm{Out}_\mathcal{V}$ is the verifier's final decision bit; and
4. B is the maximum allowed distance between $\mathcal{V}$ and $\mathcal{P}$.

4.3 Timing constraint

Let c be the speed of light (not to be confused with the challenge bits c_i). During the fast phase, $\mathcal{V}$ rejects whenever the measured round-trip time exceeds

$$\Delta t_{\max} \;:=\; \frac{2B}{c}.$$

We take double the distance B to account for the round-trip time.

4.4 Noise model

We model physical noise by assuming that each transmitted qubit, on both the challenge hop ($\mathcal{V} \to \mathcal{P}$) and the response hop ($\mathcal{P} \to \mathcal{V}$), undergoes an *independent and identically distributed* (*i.i.d.*) depolarizing channel [16] $\mathcal{D}_\eta$, so there are no correlations across the protocol rounds or across the hops. The depolarizing channel with parameter $\eta \in [0, 1]$ acts on any single-qubit projector Π as

$$\mathcal{D}_\eta(\Pi) = (1 - \eta)\Pi + \eta\, \mathbb{I}/2,$$

i.e., with probability $1 - \eta$ the qubit is unchanged, and with probability η it is replaced by the *maximally mixed* state $\mathbb{I}/2$ (complete randomness), which yields a *fair coin flip* under measurement in *any* basis.

When the parties measure in the intended bases, a single noisy hop yields the correct bit with probability

$$1 - \eta/2 = (1 - \eta) \cdot 1 + \eta/2.$$

Across a full fast round (two hops), the final bit equals the original iff either both hops preserve the bit or both flip it, so the honest per-round acceptance is

$$p(\eta) = (1 - \eta/2)^2 + (\eta/2)^2 = 1 - \eta + \eta^2/2.$$

4.5 Completeness

Completeness is the *liveness* condition: if the prover is honest and physically within the distance bound, then the protocol should accept except with negligible probability. We write $(A \leftrightarrow B)$ for the execution of the interactive protocol between parties A and B. We define completeness as follows:

Definition 2 (Completeness [10])**.** If an honest prover $\mathcal{P}$ is located at distance $d \leq B$ from the verifier $\mathcal{V}$, then

$$\Pr\Big[(\mathcal{V}(x, r) \;\leftrightarrow\; \mathcal{P}(x)) \text{ accepts}\Big] \;\geq\; 1 - \mathrm{negl}(\lambda).$$

5 Threat model

Throughout this paper, the adversary $\mathcal{A}$ is QPT and controls both classical and quantum channels: it may intercept, delay, inject, drop, relay, store systems, and apply arbitrary *completely positive, trace-preserving (CPTP)* maps to any quantum data it possesses [16]. A CPTP map is any physically allowed quantum operation on a state. All actions must respect quantum mechanics and relativity (e.g., no cloning, measurement disturbance, no superluminal signalling).

5.1 Security experiments

We follow a standard game-based template. For each threat, we define an *experiment* that specifies the parties, the adversary's capabilities, and the acceptance condition; the experiment outputs 1 iff the verifier accepts. The adversary's *advantage* is the probability that the experiment outputs 1. A protocol is *secure* against that threat if every QPT adversary achieves negligible advantage as a function of the security parameter λ.

In the following, we define the three main threat classes: distance-fraud (DF), mafia-fraud (MF), and terrorist-fraud (TF).

5.1.1 Distance-fraud

Definition 3 (Distance-fraud experiment adapted from [10])**.**

1. **Setup.** Verifier $\mathcal{V}$ and dishonest prover $\mathcal{P}^\star$ share a long-term secret key x. $\mathcal{P}^\star$ is located at distance $d > B$ from $\mathcal{V}$.
2. **Challenge session.** $\mathcal{V}$ and $\mathcal{P}^\star$ interact in a complete execution of the QDB protocol, consisting of n fast rounds obeying the distance bound limit as in an honest run.
3. **Output.** $\mathcal{V}$ outputs $\mathrm{Out}_{\mathcal{V}} \in \{0, 1\}$.

Definition 4 (Distance-fraud advantage adapted from [10])**.** For every QPT dishonest prover $\mathcal{P}^\star$ located at distance $d > B$ from verifier $\mathcal{V}$, define

$$\mathsf{Adv}^{\mathrm{DF}}_{\mathrm{QDB},\mathcal{P}^\star}(\lambda) := \Pr\left[\begin{array}{l} x \leftarrow \mathsf{KeyGen}(1^\lambda); \\ \mathrm{Out}_{\mathcal{V}} \leftarrow (\mathcal{V}(x, r) \ \leftrightarrow \ \mathcal{P}^\star(x)) \end{array} \ : \ \mathrm{Out}_{\mathcal{V}} = 1\right].$$

Definition 5 (Distance-fraud security adapted from [10])**.** A QDB protocol is $\varepsilon_{\mathrm{DF}}(\lambda)$-distance-fraud secure if, for every QPT adversary $\mathcal{P}^\star$,

$$\mathsf{Adv}^{\mathrm{DF}}_{\mathrm{QDB},\mathcal{P}^\star}(\lambda) \ \leq \ \varepsilon_{\mathrm{DF}}(\lambda),$$

where $\varepsilon_{\mathrm{DF}}$ is a negligible function.

5.1.2 Mafia-fraud

Definition 6 (Mafia-fraud experiment adapted from [10])**.**

1. **Setup.** Verifier $\mathcal{V}$ and honest prover $\mathcal{P}$ share a long-term secret key x. Two QPT adversaries, $\mathcal{A}_1$ (co-located with $\mathcal{V}$) and $\mathcal{A}_2$ (co-located with $\mathcal{P}$), share an authenticated classical and quantum channel.
2. **Learning phase.** The adversaries may initiate and control any polynomial number of auxiliary executions of the QDB protocol between $\mathcal{V}$ and $\mathcal{P}$: in each such execution, every classical and quantum message between the honest parties passes through $(\mathcal{A}_1, \mathcal{A}_2)$, who may relay, delay, drop, modify, or inject messages arbitrarily (subject only to the timing constraints). At the end of this phase, $(\mathcal{A}_1, \mathcal{A}_2)$ retain the entire classical transcript and their joint quantum state.
3. **Challenge session.** The adversaries $(\mathcal{A}_1, \mathcal{A}_2)$ interact with $\mathcal{V}$ in a full execution of the QDB protocol. Simultaneously, they may interact with the honest $\mathcal{P}$ (who uses the correct long-term key x and is at distance $d > B$). $\mathcal{V}$ enforces the distance bound B.
4. **Output.** $\mathcal{V}$ outputs $\mathrm{Out}_{\mathcal{V}} \in \{0, 1\}$.

Definition 7 (Mafia-fraud advantage adapted from [10])**.** Let $(\mathcal{A}_1, \mathcal{A}_2)$ be any QPT pair with $\mathcal{A}_1$ close to verifier $\mathcal{V}$ and $\mathcal{A}_2$ close to prover $\mathcal{P}$ at distance $d > B$ from $\mathcal{V}$. Set

$$\mathsf{Adv}^{\mathrm{MF}}_{\mathrm{QDB},\mathcal{A}_1,\mathcal{A}_2}(\lambda) := \Pr\left[\begin{array}{ll} x \leftarrow \mathsf{KeyGen}(1^\lambda); & \\ (\mathcal{A}_1, \mathcal{A}_2) \ \leftrightarrow \ \mathcal{P}(x); & : \ \mathrm{Out}_{\mathcal{V}} = 1 \\ \mathrm{Out}_{\mathcal{V}} \leftarrow (\ \mathcal{V}(x, r) \ \leftrightarrow \ \mathcal{A}_1 \parallel \mathcal{A}_2\ ) & \end{array}\right].$$

Definition 8 (Mafia-fraud security adapted from [10])**.** A QDB protocol is $\varepsilon_{\mathrm{MF}}(\lambda)$-mafia-fraud secure if, for every QPT pair $(\mathcal{A}_1, \mathcal{A}_2)$,

$$\mathsf{Adv}^{\mathrm{MF}}_{\mathrm{QDB},\mathcal{A}_1,\mathcal{A}_2}(\lambda) \ \leq \ \varepsilon_{\mathrm{MF}}(\lambda),$$

where $\varepsilon_{\mathrm{MF}}$ is a negligible function.

5.1.3 Terrorist-fraud

Terrorist-fraud models the strongest collusion attack in distance-bounding: a dishonest prover $\mathcal{P}^\star$ located outside the bound B collaborates with a nearby helper $\mathcal{A}$ co-located with $\mathcal{V}$. The goal is for $\mathcal{A}$ to convince $\mathcal{V}$ during the *fast phase*, even though $\mathcal{P}^\star$ is too far away to respond in time. In contrast to mafia-fraud, the prover itself is malicious and may deliberately provide information or quantum systems to the helper in order to pass the challenge session.

Intuitively, terrorist-fraud lies between mafia-fraud and outright key disclosure. As in mafia-fraud, a nearby helper $\mathcal{A}$ must answer the verifier during the fast phase; unlike mafia-fraud, however, the distant prover $\mathcal{P}^\star$ is itself dishonest and may deliberately assist. This assistance is nevertheless restricted: it may help only for the current session and should not equip $\mathcal{A}$ to authenticate alone in future sessions. The core TF question is therefore whether $\mathcal{P}^\star$ can provide session-specific help that lets $\mathcal{A}$ answer the fast rounds, yet remains non-transferable in the sense formalised below.

Definition 9 (Terrorist-fraud experiment adapted from [10])**.**

1. **Setup.** Verifier $\mathcal{V}$ and dishonest prover $\mathcal{P}^\star$ share a long-term secret key x. $\mathcal{P}^\star$ is located at distance $d > B$ from $\mathcal{V}$. A helper $\mathcal{A}$ is co-located with $\mathcal{V}$. The pair $(\mathcal{P}^\star, \mathcal{A})$ share an authenticated classical and quantum channel.

2. **Learning phase.** $(\mathcal{P}^\star, \mathcal{A})$ may engage in any polynomial number of auxiliary executions of the QDB protocol with $\mathcal{V}$. In each such execution, $\mathcal{A}$ is co-located with $\mathcal{V}$ and may relay, modify, or inject messages, while $\mathcal{P}^\star$ participates at its true distant location using key x. The pair may also exchange arbitrary classical and quantum messages over their channel, and all information gathered in this phase is available later.

3. **Challenge session.** $\mathcal{A}$ impersonates $\mathcal{P}$ in a full execution of the QDB protocol with $\mathcal{V}$, subject to the same distance bound B as in an honest run.

4. **Output.** $\mathcal{V}$ outputs $\mathrm{Out}_\mathcal{V} \in \{0, 1\}$.

Without further restriction, Definition 9 would be trivial: $\mathcal{P}^\star$ could simply reveal reusable secret information (e.g., the long-term key x), after which $\mathcal{A}$ could impersonate $\mathcal{P}$ not only in the current challenge session but also in future sessions. The distinguishing feature of terrorist-fraud is therefore that the prover's help is allowed to be useful for the current session, yet should not give the helper a reusable ability to authenticate on its own later. We formalise this *non-transferability* requirement as follows:

Definition 10 (Non-transferable assistance)**.** Let $(\mathcal{P}^\star, \mathcal{A})$ be a QPT pair as in Definition 9, and fix a long-term key x. Run the terrorist-fraud experiment once and consider the final (classical and quantum) state of $\mathcal{A}$ at the end of this execution. Now let $\mathcal{A}$, starting from this state and with no further interaction with $\mathcal{P}^\star$, engage alone in a second execution of the QDB protocol with $\mathcal{V}$, where the long-term key is still x and all nonces and verifier randomness are freshly sampled. We say that the assistance of $(\mathcal{P}^\star, \mathcal{A})$ is *non-transferable* if the probability that $\mathcal{V}$ accepts in this second execution is negligible in the security parameter λ.

Definition 11 (Terrorist-fraud advantage adapted from [10])**.** Let $(\mathcal{P}^\star, \mathcal{A})$ be any QPT pair as in Definition 9, whose assistance is non-transferable in the sense of Definition 10. The terrorist-fraud advantage of $(\mathcal{P}^\star, \mathcal{A})$ against QDB is

$$\mathsf{Adv}^{\mathrm{TF}}_{\mathrm{QDB},\mathcal{P}^\star,\mathcal{A}}(\lambda) := \Pr\left[\begin{array}{l} x \leftarrow \mathsf{KeyGen}(1^\lambda); \\ \mathrm{Out}_\mathcal{V} \leftarrow (\,\mathcal{V}(x,r) \;\leftrightarrow\; (\mathcal{P}^\star(x), \mathcal{A})\,) \end{array} \;:\; \mathrm{Out}_\mathcal{V} = 1\right],$$

where the interaction $(\mathcal{V} \leftrightarrow (\mathcal{P}^\star, \mathcal{A}))$ is the terrorist-fraud experiment of Definition 9.

Definition 12 (Terrorist-fraud security adapted from [10])**.** A QDB protocol is $\varepsilon_{\mathrm{TF}}(\lambda)$-terrorist-fraud secure if, for every QPT pair $(\mathcal{P}^\star, \mathcal{A})$ whose assistance is non-transferable as in Definition 10,

$$\mathsf{Adv}^{\mathrm{TF}}_{\mathrm{QDB},\mathcal{P}^\star,\mathcal{A}}(\lambda) \;\leq\; \varepsilon_{\mathrm{TF}}(\lambda),$$

where $\varepsilon_{\mathrm{TF}}$ is a negligible function.

5.1.4 Soundness

Definition 13 (Soundness w.r.t. a set of threats adapted from [10])**.** Let $\mathcal{T} \subseteq \{\mathrm{DF}, \mathrm{MF}, \mathrm{TF}\}$. A QDB protocol is $\mathcal{T}$-sound if it is secure against all threats in $\mathcal{T}$, i.e., for each $T \in \mathcal{T}$ there exists a negligible function $\varepsilon_T(\lambda)$ such that $\mathsf{Adv}^{T}_{\mathrm{QDB}}(\lambda) \leq \varepsilon_T(\lambda)$. We call the special case $\mathcal{T} = \{\mathrm{DF}, \mathrm{MF}, \mathrm{TF}\}$ *full soundness*.

5.2 Framework instantiation across QDB protocol families

Instantiating our security framework for a QDB protocol follows the same template throughout. One identifies (i) the messages exchanged in the slow and fast phases; (ii) the *per-round acceptance condition*, which always decomposes into a timing check and a protocol-specific *value or correlation check*; and (iii) a noise model that determines the honest per-round acceptance probability. Once a per-round success bound p is known for a given fraud experiment (DF/MF/TF), the same lifting argument via Lemma 1 yields the corresponding multi-round bound.

Prepare-and-measure single-qubit QDB (Abidin et al. [5], Abidin [6]). These QDB protocols fit our security framework essentially verbatim. The slow phase derives per-round secrets, each fast round uses a single-qubit challenge and a single-qubit response, and a round is accepted iff the response is timely and the verifier's measurement outcome matches the verifier's fresh challenge bit. This directly matches the DF/MF/TF experiments of Section 5.1 and yields explicit per-round cheating probabilities. The MF bound of Abidin [6] was subsequently tightened by Verschoor [11].

Entanglement-based QDB (Abidin et al. [7], Bogner et al. [8]). Entanglement-based QDB protocols replace the prepare-and-measure fast phase by operations on shared entangled systems. In our security framework, the instantiation step is to define the per-round acceptance condition as the conjunction of the timing check and a protocol-specific *correlation check*. The DF/MF/TF experiments of Section 5.1 remain unchanged. However, Bogner et al. [8] provide only an informal security analysis and do not report per-round cheating probabilities under the standard fraud experiments. Completeness analysis for these QDB protocols typically requires a noise model that captures entanglement decoherence rather than single-qubit depolarisation.

Continuous-variable QDB (Bogner et al. [9]). Continuous-variable proposals use quantum states and measurements with continuous outcomes in the fast phase. They still instantiate the experiments of Section 5.1 by choosing a per-round acceptance condition consisting of the timing check and a value check, but the latter usually takes the form of a threshold test on a continuous measurement outcome. As a result, reporting a single per-round DF/MF/TF success probability requires specifying the exact thresholding rule and the corresponding noise model. In contrast to the single-qubit setting, noise in continuous-variable protocols is often dominated by loss and Gaussian noise rather than depolarisation.

Summary. Across all QDB protocol families, the protocol-independent parts of the security framework, namely the threat experiments (Section 5.1), the timing model (Section 4.3), and the multi-round lifting (Lemma 1), remain unchanged; what varies across proposals is the definition of the per-round value predicate and the noise model needed to quantify completeness and fraud success rates.

6 Security analysis of Abidin's QDB protocol

In this section we apply the framework to the QDB protocol of Abidin [6] (Section 6.1). After recalling the protocol, we prove completeness (Definition 2), derive distance-fraud (Definition 5) and mafia-fraud (Definition 8) bounds, and analyse terrorist-fraud (Definition 12), highlighting an insecurity. We then establish two-fraud soundness (DF/MF). Each analysis first bounds single-round acceptance and is lifted to the multi-round setting via Lemma 1. The noise model (Section 4.4) is used for the completeness analysis and threshold sizing; the DF/MF bounds are noise-free. This is a conservative choice, as we assume a perfect adversary (with noiseless channels and measurements), so any additional physical noise can only make DF/MF attacks less effective.

6.1 Protocol description

Figure 1 depicts one session of the QDB protocol proposed by Abidin [6] between a verifier $\mathcal{V}$ and a prover $\mathcal{P}$. These two parties share a long-term key x. The session consists of an authenticated and public *slow* phase, where nonces and per-round secrets are set up, followed by a *fast* phase of n rapid quantum challenge-response rounds whose round-trip times are measured by $\mathcal{V}$. Finally, in the decision phase, $\mathcal{V}$ checks if the response of $\mathcal{P}$ is consistent with the challenge, if the round-trip time is within the allowed bound, and accepts or rejects the session.

6.1.1 Slow phase (setup; untimed)

1. *Nonce exchange.* $\mathcal{V}$ samples $N_v \xleftarrow{\$} \{0,1\}^n$ and sends it to $\mathcal{P}$; $\mathcal{P}$ samples $N_p \xleftarrow{\$} \{0,1\}^n$ and sends it back to $\mathcal{V}$ (straight arrows in Figure 1 are classical messages). This step is performed to ensure *freshness* of the protocol's execution.

2. *Deriving per-round secrets.* Using a keyed quantum-secure pseudorandom function (PRF) f_x [17], and the exchanged nonces, allows both parties to compute a pseudorandom $2n$-bit string $a \parallel b$ (computationally indistinguishable from uniform to any QPT adversary without knowledge of x), and divide it into two n-bit sequences

$$a \parallel b \;=\; f_x(N_v, N_p), \qquad a = (a_1, \ldots, a_n),\; b = (b_1, \ldots, b_n).$$

   The bit a_i selects the *challenge basis* for round i, while b_i selects the *response basis* for round i.

6.1.2 Fast phase (distance measurement; timed)

For each round $i = 1, \ldots, n$:

1. *Challenge preparation.* $\mathcal{V}$ samples a uniform challenge bit $c_i \xleftarrow{\$} \{0,1\}$ and prepares the challenge state in the challenge basis a_i (as defined in Section 3.1)

$$|\psi_i\rangle = |c_i\rangle_{a_i},$$

   and sends it to $\mathcal{P}$ (wavy arrow denotes a single-qubit transmission in Figure 1). $\mathcal{V}$ records the sending time as t_i^{send}.

2. *Immediate measurement.* Upon receipt, $\mathcal{P}$ measures the challenge state $|\psi_i\rangle$ in basis a_i, obtaining the measurement result c_i'. In a noiseless setting, $c_i' = c_i$.

3. *Response preparation.* $\mathcal{P}$ prepares the response state $|\phi_i\rangle$ by preparing the measurement result c_i' using the response basis b_i and sends (as defined in Section 3.1)

$$|\phi_i\rangle = |c_i'\rangle_{b_i},$$

   back to $\mathcal{V}$ (second wavy arrow in Figure 1).

4. *Response measurement.* Upon receipt, $\mathcal{V}$ measures the response state $|\phi_i\rangle$ in basis b_i, obtaining the measurement result c_i''. $\mathcal{V}$ records the receiving time as t_i^{recv}.

6.1.3 Decision phase

After n rounds, $\mathcal{V}$ checks for each round i if their measurement result is consistent with the challenge $c_i = c_i''$, and checks if the round-trip time $t_i^{\text{recv}} - t_i^{\text{send}} \leq \Delta t_{\max}$ is within the allowed bound. If both conditions are met, $\mathcal{V}$ accepts the round i. If the number of accepted rounds is at least the threshold τ, $\mathcal{V}$ accepts the session. Otherwise, $\mathcal{V}$ rejects the session.

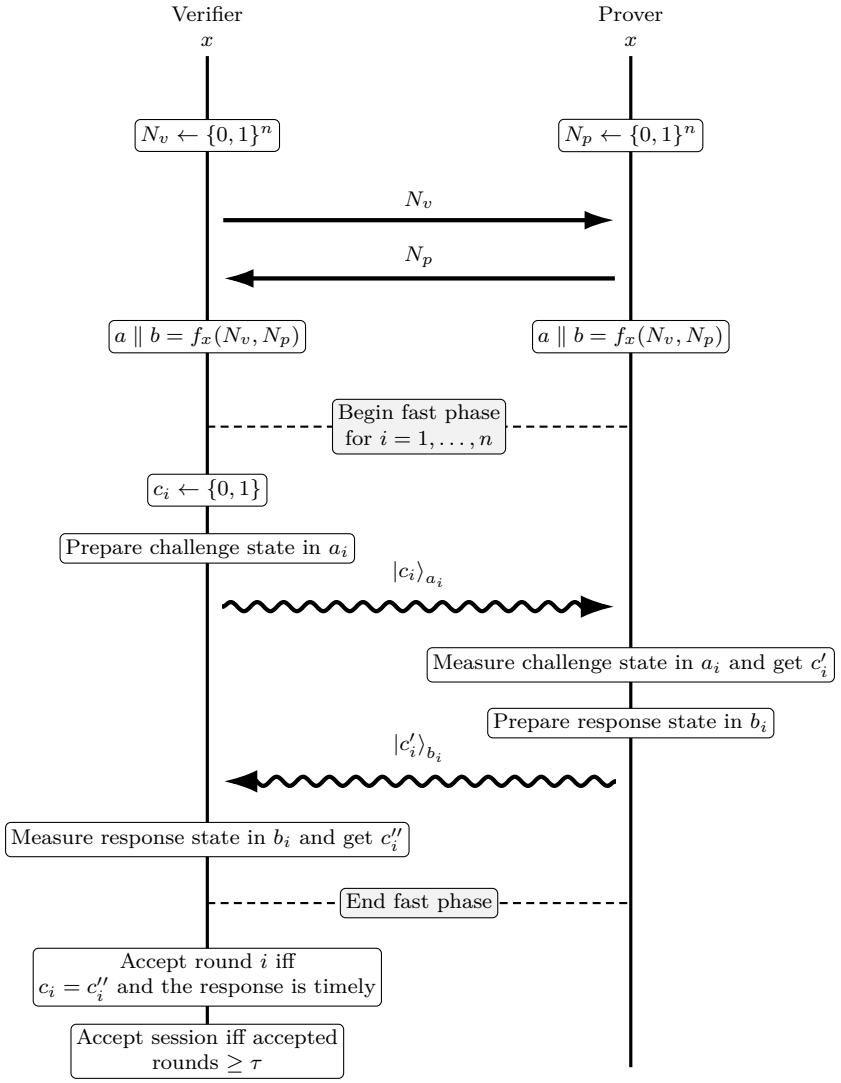


Figure 1: QDB Protocol [6], based on the classical DB protocol of Hancke-Kuhn [2].

6.2 Completeness

We prove completeness (Definition 2) for the QDB protocol of Section 6.1. A fast round is accepted iff the response is timely and the value check passes (i.e., $c_i = c_i''$). In our analysis we separate these two aspects to keep the proofs modular. Let $I_i \in \{0,1\}$ indicate acceptance in round i and $S = \sum_{i=1}^{n} I_i$ the total number of accepted rounds. The verifier accepts the session iff $S \geq \tau$, where τ is the protocol's threshold.

Lemma 2 (Single round completeness)**.** *In the ideal (noiseless) model an honest prover passes every round with probability* 1*:*

$$\Pr[I_i = 1] = 1$$

for all $i \in [n]$*.*

Proof. Timing.
Since $d \leq B$, the round-trip time is at most $2B/c = \Delta t_{\max}$, which is less than the response deadline and therefore the timing bound holds.
Value.
Let $|\psi_i\rangle = |c_i\rangle_{a_i}$ be the challenge state and $\Pi_{a_i,c_i} = |c_i\rangle_{a_i} \langle c_i|_{a_i}$ (as in Section 3.1).

$$\Pr[c_i' = c_i] = \left\| \Pi_{a_i,c_i} |\psi_i\rangle \right\|^2 = \left\| |\psi_i\rangle \right\|^2 = 1,$$

With $|\phi_i\rangle = |c_i'\rangle_{b_i}$ and $\Pi_{b_i,c_i'} = |c_i'\rangle_{b_i} \langle c_i'|_{b_i}$, the same reasoning yields

$$\Pr[c_i'' = c_i'] = \left\| \Pi_{b_i,c_i'} |\phi_i\rangle \right\|^2 = \left\| |\phi_i\rangle \right\|^2 = 1,$$

so $c_i'' = c_i' = c_i$ with probability 1, and the value test passes. Hence, $I_i = 1$. □

In the following, we prove completeness under the noise model (Section 4.4) in a multi-round setting (Lemma 1).

Theorem 3 (Completeness under i.i.d. depolarizing noise)**.** *Let each transmitted qubit pass independently through* $\mathcal{D}_\eta$ *on the forward and return channels. For any threshold* $\tau \in \{0, \ldots, n\}$ *with* $\tau < np(\eta)$*, where* $p(\eta)$ *is as in Section 4.4,*

$$\Pr[S < \tau] \;\leq\; \exp\Big(-n \cdot D\big(\tfrac{\tau}{n} \,\|\, p(\eta)\big)\Big).$$

In particular, for $\eta = 0$ *we have* $\Pr[S < \tau] = 0$ *by Lemma 2.*

Proof. Since $d \leq B$, the timing bound holds deterministically for an honest prover (Lemma 2); thus I_i reduces to the value test in each round. By the noise model (Section 4.4), the honest per-round acceptance probability equals $p(\eta)$, hence $\mathbb{E}[I_i \mid \mathcal{F}_{i-1}] = p(\eta)$ for all i. Applying Lemma 1 (lower tail) with $p = p(\eta)$ yields the claim. In the i.i.d. setting, this coincides with the binomial lower-tail bound for $S \sim \mathrm{Bin}(n, p(\eta))$. □

6.3 Distance-fraud security

Recall the distance-fraud experiment from Definition 3. The per-round secrets $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ are fixed in the slow phase and therefore *known* to $\mathcal{P}^\star$ who holds the shared secret key x. Throughout the fast phase the distant prover $\mathcal{P}^\star$ is located at distance $d > B$. Since $\mathcal{V}$ accepts a round only if it receives a response within $\Delta t_{\max} = 2B/c$ of sending the challenge, any response system received by the deadline from distance $d > B$ must have been emitted before a light-speed signal carrying $|\psi_i\rangle$ could reach the emitter. Thus, any such timely response is independent of the fresh challenge bit c_i. This is formalized in the following Lemma, which is adapted from [10].

Lemma 4 (Out-of-bound independence adapted from [10])**.** *Fix a fast-phase round* i*. Let* Near *denote the ball of radius* B *around the verifier* $\mathcal{V}$ *and let* Far *be its complement. Suppose* $\mathcal{V}$ *sends the challenge state* $|\psi_i\rangle = |c_i\rangle_{a_i}$ *at time* t_i^{send} *and accepts the round only if it receives a response system by time* $t_i^{send} + \Delta t_{\max}$*.*

Then, conditioned on the entire history available at time t_i^{send} *(and on any internal quantum state), any (classical or quantum) message that is generated entirely in* Far *and is received by* $\mathcal{V}$ *by the deadline* $t_i^{send} + \Delta t_{\max}$ *is statistically independent of the fresh challenge bit* c_i*.*

Proof. Let the message be emitted from some location at distance $\delta > B$ from $\mathcal{V}$. To arrive by $t_i^{\text{send}} + \Delta t_{\max} = t_i^{\text{send}} + 2B/c$, it must be emitted no later than $t_i^{\text{send}} + 2B/c - \delta/c < t_i^{\text{send}} + \delta/c$, i.e., strictly before any light-speed signal carrying $|\psi_i\rangle$ could reach the emitter. Therefore the emitted message cannot depend on c_i. □

Lemma 4 is a direct consequence of relativity (no superluminal signalling). In the following, we prove the single-round distance-fraud bound and then aggregate across rounds using Lemma 1.

Lemma 5 (Distance-fraud single-round bound)**.** *For every round i and any prior transcript, a distant prover's timing-respecting strategy satisfies*

$$\Pr[I_i = 1] \;\le\; \tfrac{1}{2}.$$

Proof. Timing.
Since $\mathcal{P}^\star$ is at distance $d > B$, any message from $\mathcal{P}^\star$ that reaches $\mathcal{V}$ by the deadline cannot depend on the fresh challenge bit c_i (Lemma 4). Thus, the prover's returned state $|\phi_i\rangle$ is independent of c_i and fixed by the private randomness r_i of $\mathcal{P}^\star$. This is the only time-respecting strategy that can be applied by $\mathcal{P}^\star$.
Value.
$\mathcal{V}$ measures the response in the known basis b_i and accepts the value check iff the outcome equals the uniform challenge bit c_i, which is hidden from $\mathcal{P}^\star$. Writing $\Pi_{b_i,r_i} := |r_i\rangle_{b_i}\langle r_i|_{b_i}$, we have

$$\begin{aligned}\Pr[I_i = 1 \mid |\phi_i\rangle] &= \tfrac{1}{2} \sum_{r_i \in \{0,1\}} \langle\phi_i|\Pi_{b_i,r_i}|\phi_i\rangle \\ &= \tfrac{1}{2}\,\langle\phi_i|(\Pi_{b_i,0} + \Pi_{b_i,1})|\phi_i\rangle \\ &= \tfrac{1}{2}\,\langle\phi_i|\phi_i\rangle \\ &= \tfrac{1}{2},\end{aligned}$$

since $\Pi_{b_i,0} + \Pi_{b_i,1} = \mathbb{I}$. Averaging over $|\phi_i\rangle$ gives $\Pr[I_i = 1] \le 1/2$. Sending a response after the deadline can only decrease the success probability. □

For the multi-round distance-fraud security, we have the following Theorem.

Theorem 6 (Distance-fraud multi-round bound)**.** *For any timing-respecting strategy and any threshold τ with $\tau > n/2$,*

$$\Pr[S \ge \tau] \;\le\; \exp\Big(-\,n \cdot D\big(\tfrac{\tau}{n} \,\|\, \tfrac{1}{2}\big)\Big).$$

Proof. By Lemma 5 we have $\mathbb{E}[I_i \mid \mathcal{F}_{i-1}] \le 1/2$ for the natural filtration capturing the transcript and the adversary's state. Applying Lemma 1 with $p = 1/2$ therefore yields the stated bound for every threshold $\tau > n/2$. □

6.4 Mafia-fraud security

Recall the mafia-fraud experiment defined in Definition 6. The per-round basis strings $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_n)$ are derived from the long-term key x and the public nonces (N_v, N_p) via a quantum-secure PRF. Consequently, they remain computationally indistinguishable from uniform strings for any QPT adversary pair $(\mathcal{A}_1, \mathcal{A}_2)$ lacking knowledge of x. In the fast phase, only the adversary $\mathcal{A}_1$ (located near the verifier) can respond within the time bound B; $\mathcal{A}_2$ may assist only through the (untimed) learning phase.

Abidin's original analysis of this protocol estimated the single-round mafia-fraud success probability at 3/4 [6]. However, this underestimates the adversary's advantage. Verschoor later demonstrated that by combining a more capable pre-ask phase with a tailored response strategy,

a mafia adversary can achieve a success probability of 7/8 per round [11]. We summarize Verschoor's attack using our notation below. The core observation is that an adversary can correctly respond to the verifier's challenge simply by knowing whether a_i and b_i belong to the same basis. For each fast round i, define the parity bit:

$$k_i := a_i \oplus b_i \in \{0, 1\}.$$

The verifier's challenge state is $|\psi_i\rangle = |c_i\rangle_{a_i}$, and an honest prover's response is $|\phi_i\rangle = |c_i\rangle_{b_i}$. The honest prover's operation can be characterised as:

- if $k_i = 0$ (same basis): applying the identity gate (reflecting the challenge), and
- if $k_i = 1$ (different bases): applying the Hadamard gate (mapping the basis $\{|0\rangle_0, |1\rangle_0\}$ to $\{|0\rangle_1, |1\rangle_1\}$ and vice versa).

Thus, knowledge of k_i suffices for an adversary near the verifier to emulate the honest prover's behaviour during the fast phase.

Verschoor's attack is an intra-session *pre-ask* strategy. It exploits the time gap between the slow phase and the fast phase within the *challenge session*.

1. *Slow phase relay.* $(\mathcal{A}_1, \mathcal{A}_2)$ relay the slow phase messages between $\mathcal{V}$ and $\mathcal{P}$. Consequently, all parties agree on the fresh nonces N_v, N_p and derive the same session-specific secrets a, b.
2. *Pre-ask (Extraction).* Before the verifier starts the fast phase, $\mathcal{A}_2$ (near $\mathcal{P}$) initiates a fast phase execution with $\mathcal{P}$. For each round i, $\mathcal{A}_2$ sends the intermediate state $|\xi\rangle = \cos(\frac{3\pi}{8})|0\rangle + \sin(\frac{3\pi}{8})|1\rangle$. The honest $\mathcal{P}$ measures in basis a_i and returns the result. $\mathcal{A}_2$ measures this response to obtain a guess k_i' for the parity $k_i = a_i \oplus b_i$, which is correct with probability 3/4 [11].
3. *Fast phase challenge.* When $\mathcal{V}$ sends the actual challenge $|c_i\rangle_{a_i}$ to $\mathcal{A}_1$, $\mathcal{A}_1$ uses the pre-computed guess k_i'. If $k_i' = 1$, they apply a Hadamard gate (basis swap) to the challenge; otherwise, they reflect it.

Lemma 7 (Mafia-fraud single-round attack [11])**.** *There exists a mafia-fraud adversary pair $(\mathcal{A}_1, \mathcal{A}_2)$ in the experiment of Definition 6 such that, in each fast round i, the acceptance indicator I_i satisfies*

$$\Pr[I_i = 1] \;=\; \tfrac{3}{4} \cdot 1 \;+\; \tfrac{1}{4} \cdot \tfrac{1}{2} \;=\; \tfrac{7}{8}.$$

Proof. Timing.
By definition of the mafia-fraud experiment (Definition 6), $\mathcal{A}_1$ is located near $\mathcal{V}$ during the fast phase and can therefore return a response within the deadline.
Value.
Based on the pre-ask analysis above, $\Pr[k_i' = k_i] = 3/4$ and $\Pr[k_i' \neq k_i] = 1/4$. If $k_i' = k_i$, $\mathcal{A}_1$'s transformation produces exactly $|c_i\rangle_{b_i}$, so the verifier accepts with probability 1. If $k_i' \neq k_i$, the verifier measures in the wrong basis, and the outcome equals c_i with probability 1/2. Conditioning on these two cases yields:

$$\Pr[I_i = 1] = \Pr[k_i' = k_i] \cdot 1 + \Pr[k_i' \neq k_i] \cdot \tfrac{1}{2} = \tfrac{3}{4} + \tfrac{1}{4} \cdot \tfrac{1}{2} = \tfrac{7}{8}.$$

□

We establish the multi-round security bound in the following.

Theorem 8 (Mafia-fraud multi-round bound for Verschoor's attack)**.** *Let $(\mathcal{A}_1, \mathcal{A}_2)$ be the mafia-fraud adversary pair from Lemma 7, and let $S = \sum_{i=1}^{n} I_i$ be the number of accepted rounds when they execute that strategy independently in each of the n fast rounds. Then, for any threshold τ with $\tau > \frac{7}{8}n$,*

$$\Pr[S \geq \tau] \;\leq\; \exp\Big(-\,n \cdot D\big(\tfrac{\tau}{n} \,\|\, \tfrac{7}{8}\big)\Big).$$

Proof. For this fixed adversary pair $(\mathcal{A}_1, \mathcal{A}_2)$, Lemma 7 gives $\mathbb{E}[I_i \mid \mathcal{F}_{i-1}] = 7/8$ for every round i, where $\mathcal{F}_{i-1}$ is the natural filtration capturing the transcript and the adversary's internal state. Applying Lemma 1 with $p = 7/8$ therefore yields the stated bound on $\Pr[S \geq \tau]$ for all $\tau > \frac{7}{8}n$. □

6.5 Terrorist-fraud security

Recall the terrorist-fraud experiment Definition 9, where a distant dishonest prover $\mathcal{P}^\star$ collaborates with a nearby helper $\mathcal{A}$ who obeys the distance bound B.

The QDB protocol [6] is inherently insecure against TF because the fast-phase behaviour of an honest prover is fully determined by the session-specific basis strings (a, b) that are computed in the untimed slow phase. Concretely, once a party knows (a_i, b_i) for a given round i, it can perfectly replicate the honest prover's fast-phase behaviour by measuring the incoming challenge state in basis a_i to recover the encoded bit, and immediately re-encoding that same bit in basis b_i and sending it back. No additional secret information (beyond the bases) is needed during the fast phase.

In the terrorist-fraud experiment Definition 9, the prover itself is dishonest and holds the long-term key x. After the public nonces (N_v, N_p) are fixed in the slow phase, $\mathcal{P}^\star$ can compute the session string $(a, b) = f_x(N_v, N_p)$ and transmit it to the nearby helper $\mathcal{A}$ before the fast phase begins. During the fast phase, $\mathcal{A}$ is co-located with $\mathcal{V}$ and therefore satisfies the timing bound. Moreover, for each fast round i, upon receiving the challenge $|c_i\rangle_{a_i}$, the helper $\mathcal{A}$ measures in basis a_i and obtains c_i', which equals c_i in the noiseless model, and then prepares and returns $|c_i'\rangle_{b_i}$. This is exactly the honest response behaviour described in Section 6.1, so $\mathcal{V}$ accepts each round with probability 1 by Lemma 2 in the noiseless setting.

Crucially, the information shared by $\mathcal{P}^\star$ with $\mathcal{A}$ can remain *non-transferable* in the sense of Definition 10. The strings (a, b) are tied to the current session via the fresh nonces and are not reusable across further sessions. In a fresh execution with new nonces (N_v', N_p'), the required bases are $(a', b') = f_x(N_v', N_p')$, which cannot be computed from (a, b) without knowing x. Under the quantum-secure PRF assumption, revealing (a, b) for one nonce pair does not expose the long-term key x nor enable computing the bases for fresh nonces, so $\mathcal{A}$ alone cannot (except with negligible probability) make $\mathcal{V}$ accept in a subsequent execution with freshly sampled nonces. In other words, the protocol does not enforce that any help sufficient to pass the *fast* rounds must also be transferable. This is precisely the structural feature that makes perfect TF collusion possible under Definition 9.

Hence there exists a pair $(\mathcal{P}^\star, \mathcal{A})$ with non-transferable assistance whose terrorist-fraud advantage is non-negligible, so the QDB protocol is *not* terrorist-fraud secure in the sense of Definition 12.

6.6 Soundness

Recall the definition of soundness from Definition 13. Based on the single-round security analyses of the QDB protocol, we establish the following bounds:

- An upper bound of 1/2 on the per-round distance-fraud success probability (Lemma 5);
- A mafia-fraud strategy that succeeds with per-round probability 7/8 (Lemma 7); this is currently the best known attack, but we do not know if it is optimal;

- The protocol is not terrorist-fraud secure under our model (see Section 6.5).

Consequently, we restrict our focus to $\{DF, MF\}$-soundness. While the strategy detailed in Lemma 7 represents the most effective known mafia-fraud attack against this protocol, neither we nor Verschoor [11] claim that this attack is optimal. To ensure a negligible false-acceptance rate against both distance-fraud and mafia-fraud simultaneously, the acceptance threshold must be set strictly above the maximum per-round cheating probability. We define this threshold using a tunable slack parameter ε'.

Theorem 9 (Soundness given per-round cheating bounds)**.** *Let $p_{\mathrm{DF}}^{\max} \leq 1/2$ denote the upper bound on the per-round distance-fraud success probability (where $p_{\mathrm{DF}}^{\max} = 1/2$ by Lemma 5), and let $p_{\mathrm{MF}}^{\max} < 1$ be any upper bound on the per-round mafia-fraud success probability for the adversary class under consideration. Fix any $\varepsilon' > 0$ such that $\max(p_{\mathrm{DF}}^{\max}, p_{\mathrm{MF}}^{\max}) + \varepsilon' < 1$ and set the (integer) acceptance threshold*

$$\tau := \left\lceil n\big(\max(p_{\mathrm{DF}}^{\max}, p_{\mathrm{MF}}^{\max}) + \varepsilon'\big) \right\rceil .$$

Then, letting $S = \sum_{i=1}^{n} I_i$ denote the number of accepted rounds:

$$\varepsilon_{\mathrm{DF}} = \Pr[S \geq \tau \mid DF] \;\leq\; \exp\Big(-n \cdot D\big(\tfrac{\tau}{n} \,\|\, p_{\mathrm{DF}}^{\max}\big)\Big),$$

$$\varepsilon_{\mathrm{MF}} = \Pr[S \geq \tau \mid MF] \;\leq\; \exp\Big(-n \cdot D\big(\tfrac{\tau}{n} \,\|\, p_{\mathrm{MF}}^{\max}\big)\Big).$$

In particular, if $n = \mathrm{poly}(\lambda)$ and $p_{\mathrm{MF}}^{\max} < 1$ is constant, then both $\varepsilon_{\mathrm{DF}}$ and $\varepsilon_{\mathrm{MF}}$ are negligible in λ. Since $p_{\mathrm{MF}}^{\max} \geq p_{\mathrm{DF}}^{\max}$ in our setting, it suffices in practice to determine n and τ based on $p_{\mathrm{MF}}^{\max}$.

Applying Theorem 9 to the QDB protocol, we use $p_{\mathrm{DF}}^{\max} = 1/2$ from Lemma 5. Regarding mafia-fraud, the attack by Verschoor [11] achieves a per-round success probability of 7/8 (Lemma 7). Instantiating Theorem 9 with $p_{\mathrm{MF}}^{\max} = 7/8$ yields explicit bounds on the false-accept probability against this specific strategy and any other attack with a success rate of at most 7/8. Should a more effective attack be discovered, the parameters can be adjusted by substituting the updated success probability into Lemma 1.

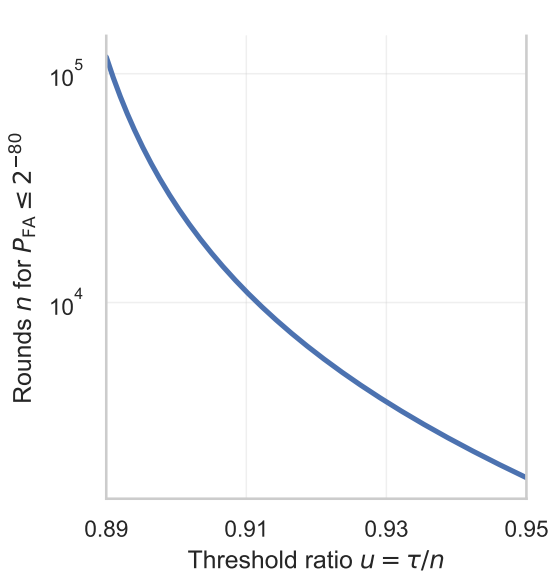
To ensure completeness in the presence of noise, the expected honest per-round acceptance probability $p(\eta)$ from Theorem 3 must strictly exceed the threshold ratio τ/n. Specifically, an honest execution is accepted with probability $1-\mathrm{negl}(\lambda)$ provided that: $p(\eta) = 1-\eta+\eta^2/2 > \tau/n$. Equivalently, for a given threshold fraction $u := \tau/n \in (\max(p_{\mathrm{DF}}^{\max}, p_{\mathrm{MF}}^{\max}), 1)$ (where $u \in (7/8, 1)$ for our instantiation), the depolarizing parameter must satisfy:

$$\eta \;<\; 1 - \sqrt{2u - 1}\,.$$

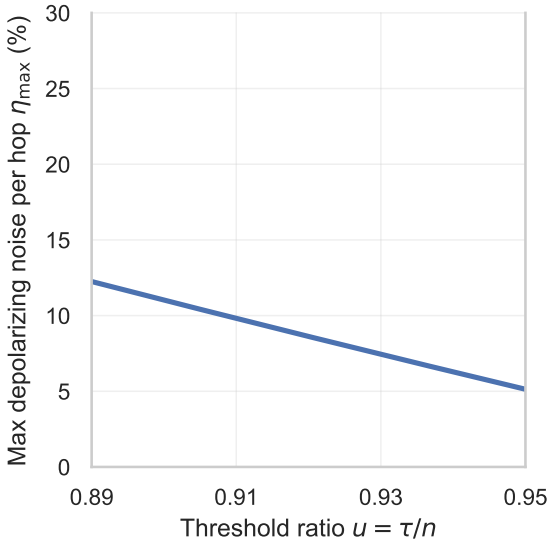
Figure 2 illustrates the performance trade-offs for the QDB protocol: (a) soundness, showing the number of rounds n required to achieve $P_{\mathrm{FA}} \leq 2^{-80}$ across varying threshold ratios $u = \tau/n$; and (b) completeness, showing the maximum tolerable depolarizing noise $\eta_{\max}$ per hop.

6.7 Towards terrorist-fraud resistance

The TF attack described in Section 6.5 succeeds because a nearby helper can reproduce the complete fast-phase behaviour given only session-specific information that the dishonest prover can disclose before the fast phase begins. These directions should be understood as structural redesign options rather than simple local patches to the protocol of [6]. Preventing such collusion therefore requires that any assistance sufficient to answer the fast rounds either (i) cannot be provided without violating the distance bound, or (ii) necessarily enables the helper to impersonate the prover in future sessions, and is thus excluded by the non-transferability requirement of Definition 10.



(a) Rounds n required for soundness ($P_{\mathrm{FA}} \leq 2^{-80}$).



(b) Maximum depolarizing noise $\eta_{\max}$ for completeness.

Figure 2: Performance trade-offs as a function of the threshold ratio $u = \tau/n$: (a) soundness; (b) completeness.

A common design direction in classical DB protocols is to bind the fast-phase responses to long-term secret material, so that producing correct responses requires access to a reusable secret (and in such protocol designs this can enable extraction). Some classical DB protocols implement this principle by coupling the responses directly to the prover's long-term key, so that enabling an accomplice to answer reliably entails disclosing enough information to recover it [18].

A complementary direction is to increase the amount of fresh challenge entropy that must be handled during the fast phase, for instance by using multi-bit or multi-state challenges, and to make the expected response depend on this larger challenge space. Intuitively, if correct fast-phase behaviour corresponds to a keyed response function over a larger challenge space, then enabling a helper without revealing the long-term secret may require providing a large amount of challenge-dependent information (in the extreme, an explicit response table) whose size grows with the challenge space, at the cost of additional fast-phase bandwidth and processing. Classical DB protocols explore this trade-off by extending binary challenges to larger challenge spaces [19].

A third, system-level mitigation is to execute the key-dependent parts of the protocol inside tamper-resistant hardware on the prover side, such that the long-term key x and any session-derived secrets, including the basis strings (a, b), are not readable or exfiltratable by the holder of the device. In this setting, the TF attack is blocked at its source because even a malicious prover cannot leak (a, b) to a nearby helper ahead of the fast phase, since it cannot access them from the secure hardware. This hardware assumption has been noted in the DB literature as a practical way to obtain robustness against TF [20].

Developing concrete protocol modifications is left for future work. Our goal is to identify the underlying cause of TF insecurity in [6] under the model of Definition 9. The fast phase can be entirely reproduced from session-derived secrets computed in the slow phase, allowing a dishonest prover to enable a nearby helper without revealing the long-term key.

7 Conclusion

In this work we proposed a reusable, game-based security framework for QDB protocols and applied it to the QDB protocol [6]. The framework adapts the classical methodology [10] to

the quantum setting by making explicit the quantum-capable adversary model and the formal distance-, mafia-, and terrorist-fraud experiments, fixing a simple noise model, and isolating the per-round success probabilities that drive the overall security guarantees. Within this setting we proved completeness of the protocol in the presence of depolarizing noise by characterising the honest per-round acceptance probability and lifting it to the multi-round setting; for active adversaries we proved an upper bound of $1/2$ on the per-round distance-fraud success probability, and showed that the best known mafia-fraud strategy succeeds with probability $7/8$ per round. Plugging these per-round bounds into our generic Lemma 1 yields explicit multi-round soundness bounds (Theorem 9) that yield negligible false-accept probability for suitable choices of n and τ. Our analysis also shows that the protocol is not terrorist-fraud secure under our model (as established in Section 6.5); we therefore treat terrorist-fraud resistance as a future research direction rather than claiming a drop-in fix.

Conceptually, the framework separates protocol-specific details (such as the encoding and message flow) from reusable security ingredients (adversary experiments, noise and timing assumptions, and concentration bounds). This makes it straightforward to compare different QDB proposals on a common basis.

We conclude by highlighting several promising directions that remain open:

- *Beyond i.i.d. depolarizing noise.* We adopt a simple, protocol-agnostic noise model in Section 4.4 as a baseline for completeness and parameter sizing. Realistic implementations of QDB protocols will also exhibit photon loss and erasures, detector inefficiency, dark counts and background clicks, basis-dependent misalignment (different error rates in Z vs. X), and potentially time-varying or correlated noise across rounds. A natural extension is to replace the simple depolarizing model with more realistic per-round noise and an explicit detection and erasure model, and derive the corresponding acceptance probabilities and protocol parameters accordingly. In particular, any treatment of *inconclusive* rounds, such as those caused by photon losses, must be modelled carefully, since an adversary could try to induce selective losses on hard rounds and answer only when confident.

- *Terrorist-fraud resistance.* Our analysis shows that the QDB protocol [6] is vulnerable to terrorist-fraud. A natural next step is to design and analyse QDB protocols that achieve full DF/MF/TF soundness within the same framework.

- *Multiple parties.* We focused on the single-verifier, single-prover setting. Extending the framework to multiple verifiers and/or multiple provers is an important direction for making QDB applicable to realistic localisation systems.

- *Experimental validation.* Validating the resulting bounds in a full experimental implementation on state-of-the-art hardware would further bridge the gap between the theoretical guarantees and real-world deployments.

Addressing these points will strengthen the case for QDB as a practical building block for secure localisation and access control, and our framework is intended to serve as a starting point for such follow-up work.

References

[1] Stefan Brands and David Chaum. Distance-Bounding Protocols. In *Workshop on the Theory and Application of Cryptographic Techniques*, pages 344–359. Springer, 1993.

[2] Gerhard P Hancke and Markus G Kuhn. An RFID Distance Bounding protocol. In *First international conference on security and privacy for emerging areas in communications networks (SECURECOMM'05)*, pages 67–73. IEEE, 2005.

[3] Gerhard P Hancke. Design of a Secure Distance-Bounding Channel for RFID. *Journal of Network and Computer Applications*, 34(3):877–887, 2011.

[4] William K Wootters and Wojciech H Zurek. A Single Quantum Cannot be Cloned. *Nature*, 299(5886):802–803, 1982.

[5] Aysajan Abidin, Eduard Marin, Dave Singelée, and Bart Preneel. Towards Quantum Distance Bounding protocols. In *International Workshop on Radio Frequency Identification: Security and Privacy Issues*, pages 151–162. Springer, 2016.

[6] Aysajan Abidin. Quantum Distance Bounding. In *Proceedings of the 12th Conference on Security and Privacy in Wireless and Mobile Networks*, pages 233–238, 2019.

[7] Aysajan Abidin, Karim Eldefrawy, and Dave Singelée. Entanglement-based Mutual Quantum Distance Bounding. In *International Symposium on Cyber Security, Cryptology, and Machine Learning*, pages 219–235. Springer, 2024.

[8] Kevin Bogner, Dave Singelée, and Aysajan Abidin. Entangled States and Bell's Inequality: A New Approach to Quantum Distance Bounding. In *2024 IEEE Symposium on Computers and Communications (ISCC)*, pages 1–6. IEEE, 2024.

[9] Kevin Bogner, Aysajan Abidin, and Dave Singelée. Continuous Variable Quantum Distance Bounding. In *IEEE INFOCOM 2025-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pages 1–6. IEEE, 2025.

[10] Ioana Boureanu, Aikaterini Mitrokotsa, and Serge Vaudenay. Practical and Provably Secure Distance-Bounding. *Journal of Computer Security*, 23(2):229–257, 2015.

[11] Sebastian Reynaldo Verschoor. *Quantum Information in Security Protocols*. PhD thesis, University of Waterloo, 2022.

[12] Aysajan Abidin. On Detecting Relay Attacks on RFID Systems using Qubits. *Cryptography*, 4(2):14, 2020.

[13] Harry Buhrman, Nishanth Chandran, Serge Fehr, Ran Gelles, Vipul Goyal, Rafail Ostrovsky, and Christian Schaffner. Position-based Quantum Cryptography: Impossibility and Constructions. *SIAM Journal on Computing*, 43(1):150–178, 2014.

[14] Charles H. Bennett and Gilles Brassard. Quantum Cryptography: Public Key Distribution and Coin Tossing. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, Bangalore, India, 1984. IEEE. Reprinted in *Theoretical Computer Science* 560 (2014), 7–11.

[15] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomization and Probabilistic Techniques in Algorithms and Data Analysis*. Cambridge University Press, Cambridge, UK, 2017.

[16] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, Cambridge, UK, 2010.

[17] Mark Zhandry. How to Construct Quantum Random Functions. *Journal of the ACM (JACM)*, 68(5):1–43, 2021.

[18] Laurent Bussard and Walid Bagga. Distance-bounding Proof of Knowledge to Avoid Real-Time Attacks. In *IFIP international information security conference*, pages 223–238. Springer, 2005.

[19] Gildas Avoine, Christian Floerkemeier, and Benjamin Martin. RFID Distance Bounding Multistate Enhancement. In *International conference on cryptology in India*, pages 290–307. Springer, 2009.

[20] Dave Singelée and Bart Preneel. Distance Bounding in Noisy Environments. In *European workshop on security in ad-hoc and sensor networks*, pages 101–115. Springer, 2007.